

A11101 295014

NAT'L INST OF STANDARDS & TECH R.I.C.



A11101295014

/Bulletin of the Bureau of Standards
QC1 .U5 V14:1918-19 C.1 NBS-PUB-C 1905







VARIANCE OF MEASURING INSTRUMENTS AND ITS RELATION TO ACCURACY AND SENSITIVITY

By Frederick J. Schlink

CONTENTS

	Page
I. Introduction: Fundamental definitions.....	741
II. Accuracy.....	741
III. Sensitivity.....	743
1. Definition.....	743
2. Passiveness distinguished from insensitiveness.....	744
IV. Variance.....	746
1. Definition.....	746
2. Types of variance.....	750
(a) Backlash.....	750
(b) Irregular variance.....	755
(c) Drift.....	757
3. Means for reducing variance.....	757
(a) Details of design.....	757
(b) Effects of vibration.....	760
4. Relation of variance to useful sensitivity.....	760
V. Summary.....	762

I. INTRODUCTION: FUNDAMENTAL DEFINITIONS

The distinction between accuracy and sensitivity in engineering and physical measuring instruments is not always carefully maintained in the discussion and calibration of such instruments. In fact certain phases of the meaning of these two terms have not been clearly expounded. The fundamental concepts are, of course, clear:

(1) An instrument is accurate when its indications accord with the true values of the quantity being measured.

(2) An instrument is sensitive when a change in the quantity being measured is accompanied by a change in the indication of the instrument.

II. ACCURACY

Neither of the above definitions is of much utility until amplified and expressed quantitatively. In the first place, instrumental accuracy as the term is defined above, is only an ideal,

since it is impossible to obtain perfect exactitude in the reading of any measuring instrument on account of divers unavoidable mechanical limitations imposed by the coaction of the working parts; the existence of friction; alterations in the properties of measuring elements with time and with changes in the environment of temperature, atmospheric pressure, humidity, etc.; and normally uncontrollable variations in the manner, frequency, and rate of operation.

In the common measurements of industry and commerce sufficient accuracy is sought to permit of dispensing with the use of corrections, allowing the employment of observed readings directly, partly to eliminate errors of calculation and partly to enable such indispensable instruments as pressure gages, measuring rules, micrometer calipers, etc., to be used by persons not possessed of special technical training. In the laboratory, on the other hand, the numerical equality of the instrumental reading to the true value of the quantity being measured is not essential (except for the arithmetical convenience which results from having small values for the corrections) since the exactness of the work will usually require that the corrections of the instrument be determined numerically at relatively close intervals throughout its reading range, unless those corrections are very small. The readings of the instrument, therefore, will be corrected before conclusions are finally drawn from them. By so doing the convenience requisite in industrial or commercial instruments is sacrificed, while exactness is served.

Instrumental accuracy is usually expressed numerically, in a negative sense, by defining the error or correction for various values of the quantity being measured. The error arising from whatever source, observed in an indication of the instrument, divided by the true value of the measured quantity may be termed the relative or *specific inaccuracy* at a given reading, the negative term being justifiable on the basis of custom and ease in application. With this in mind it is seen that *specific accuracy* may be expressed as the ratio of the value of the quantity being measured to the error of the instrumental indication at that value (this of course being the reciprocal of the quantity defined above). The ratio expressing instrumental accuracy, while not of value in the ordinary use of instruments, will be of service in rating the performance of an instrument.

III. SENSITIVITY

1. DEFINITION

By definition, any instrument which shows a change of reading for any change, however great, in the quantity being measured, is sensitive or has sensitivity. Clearly, then, this term has little significance until means are available for expressing it numerically. For particular instruments, especially those read by null methods, sensitivity has often been expressed in terms of the smallest change of value of the quantity being measured capable of causing a perceptible response or change in the reading of the instrument. Expressed in this manner, the term is necessarily unscientific and loose in its significance, since differences in the observers' acuteness of perception, or personal sensitivity, as it might be called, would result in widely varying estimates of the sensitivity of the instrument; moreover, using this criterion, we should fail to distinguish between insensitiveness and sluggishness, a distinction which, as later considerations will show, is an important one.

For example, the sensitivity of scales and balances was long defined in terms of the smallest added load which would effect a perceptible change in the position of equilibrium of the beam or pointer. The current method of expressing sensitivity in terms of measured motion at the pointer tip is an improvement, and the best current practice is to define the sensitivity of a balance as the number of linear units by which the rest point is displaced—when measured at the pointer tip—for a unit change in the load on the load pan or platform. Balance makers have not commonly adopted this practice, however, and it is still customary with them in their catalogues to refer to balances as being sensitive to a given number of milligrams when no definite value of pointer motion is meant, but merely a visually perceptible one (probably somewhat less than a millimeter). For instruments of this character, even the method of basing the sensitivity on motion at the pointer tip does not seem satisfactory, since it depends upon the accident of pointer length, and, on the basis of this definition, a given balance can be endowed with almost any desired sensitivity by the use of an optical lever or other means of high magnification.¹

A better basis would be the angular deflection of the rest position of the beam per unit addition of load to the load pan, the first

¹ The author is not unmindful of the practical limitations upon the use of a high magnification with a balance or other measuring instrument of mediocre quality. This question will be discussed later in considering the factor of variance.

expressed in radians, and the second in grams. True, this manner of expression is open to an objection from the strictly practical point of view, in that no balance, and in fact few measuring instruments commonly read by the null method, are ever subjected to a deflection of so much as a radian; the unit of angle suggested is simply larger than the practical deflection, instead of being smaller, as custom might seem to demand. All actual measurements of sensitivity, however, must for other reasons be restricted to small deflections, principally because the sensitivity itself is subject to more or less regular variation over the scale of deflections; the deflection used in determining sensitivity must be small enough, therefore, that variations over that range of deflection are quite negligible.

2. PASSIVENESS DISTINGUISHED FROM INSENSITIVENESS

Certain aspects of the action of friction in introducing error into instrumental readings have been discussed before, notably in Gramberg's *Technische Messungen* and in Weinstein's *Physikalische Maasbestimmungen*.² The variance error to be discussed in the next section has its rise in friction, in that the divagation of reference points or axes from the paths or positions determined under the assumption of a perfect and ideally fitted mechanism, is an effect which would not obtain except for the resistance which the friction offers against the taking up of the minimum or most stable positions by the journals. That action, dependent though it is upon friction, is of such definite importance from another viewpoint that it will be treated separately.

Simple frictional resistance to turning or sliding, however, has a bearing upon the ordinary concepts of sensitiveness which should be mentioned in the present connection. Its effect is to retard or delay the motion of the indicating element for both increasing and decreasing values of the quantity being measured. The amount of the static friction will be roughly proportional to the restoring force in the system, so that in the usual case, as the reading increases, there will be a tendency for proportionate increase of mechanical hysteresis due to friction, and the error which it causes will be greater, the less the forces available to effect motion of the mechanism, for a unit change in the measured quantity.

² See also the author's paper: *A Stabilized-Platform Weighing Scale of Novel Design*, Technologic Paper No. 106 of the Bureau of Standards, 1913, pp. 9-12.

The immediate effect of turning or sliding friction is to prevent response of the instrument reading to certain small changes in the measured quantity. On this account it becomes necessary to draw a distinction between sensitivity and immediacy of response. It is clear that if our first and commonly accepted definition of sensitivity be strictly adhered to, we will come upon the peculiar circumstance of an instrument having zero sensitivity whenever any finite change in the quantity being measured is accompanied by no deflection whatever of the indicating element. Nevertheless, exactly this occurs in an instrument having friction or lost motion. In fact, it is, often true in instruments of less precise character that a considerable change in the measured quantity can be effected without the occurrence of any motion whatever of the indicator; this period of inaction of the indicator persists during the taking up of slack and the overcoming of the static friction of the operating parts, much as the first pull of a locomotive on a train fails to produce motion of the caboose until the slack between all the cars has been taken up, and until the static tractive resistance of all the cars has been overcome.

With this in mind, it will be admitted that the usual definition of sensitivity leads to an absurdity, in that the determination of the sensitivity of an instrument would depend upon the absolute rather than the relative magnitudes of the quantities entering into the observation, and in that it would, moreover, involve an abrupt changing over from zero sensitivity to finite sensitivity. In view of the irrational nature of any such assumption, the definition of sensitivity must be modified, thus: Sensitivity in an instrument is the rate of change in the indication of such instrument with respect to change in the quantity being measured, it being necessarily assumed for the purposes of this definition that friction and lost motion in the mechanism have been eliminated or are negligible. (A similar postulate applies to the (analogous) determination of the scale value in instruments graduated directly in the units of the quantity being measured.) We thus distinguish between instrumental passiveness (or sluggishness) and instrumental insensitiveness, a distinction which, so far as known to the writer, has not hitherto been set forth.

The factor of passiveness may then be determined by noting the smallest alteration in the quantity to be measured which will produce any change whatever in the indication of the instrument. The amount of the least alteration in the value of the measured

quantity producing instrumental response, divided by the initial value of the measured quantity, may be called the passivity of the instrument at that point.

For the present purpose, then, we must revert to the term "perceptible movement," which has properly been rejected in so far as it concerns definitions of sensitivity. This is logical on the ground that the passiveness evidenced by delayed response has been overcome so soon as any motion whatever of the indicating element has taken place. This factor when present can be observed in the complete hysteresis loop of the instrument, as measured by the length of the horizontal line which appears in the *measured quantity—indication* curve of the instrument at points of reversal of reading. (See under Backlash, *infra*.) Passiveness then is a special case of the phenomenon of variance, the factor discussed in the section immediately following. The present section has been introduced in this order on account of the intimate relationship of passiveness to the determination of sensitivity, observations of which are peculiarly subject to error in the presence of high instrumental passivity.

IV. VARIANCE

1. DEFINITION

The third important factor to determine in the calibration of a measuring instrument is that of variance, which is defined as the range, at any given value of the measured quantity, of variation in reading which may be exhibited by the instrument under repeated application of the same value of the quantity being measured, after a steady reading has been attained, the environment remaining unchanged. This quantity, which overlaps the passiveness factor defined above, may also be called the range of uncertainty of indication, in that it represents the range within which the readings of the instrument may be expected to lie when all causes of variation save those inherent in the instrument are eliminated. The specific variance or *variancy* (the same etymological distinction being maintained as heretofore) may be defined as the ratio of the range, at any given value of the measured quantity, of variation in reading which may be exhibited by the instrument under repeated application of the same value of the quantity being measured, divided by the value of the measured quantity itself, the same assumptions applying as above as to the attainment of a steady state of indication

(see Drift, *infra*) and as to the maintenance of unchanged environment. This factor has rarely been determined in tests of measuring instruments; ignoring it in their use, as is commonly done, may cause appreciable error, and it is therefore important that it be recognized or expressed.

In the case of the usual direct-reading instrument, the variance is disclosed as the displacement observed between the upward and downward branches of the hysteresis loop when the instrument is subjected to a complete cycle of operation from a lower to a higher indication, returning again to the lower indication, while plotting point by point the instrumental readings against actual values of the quantity being measured. It is obvious from the nature of the hysteresis loop and the causes underlying it that the amount of the variance will depend upon the previous history of the instrument and, specifically, upon the immediately precedent cycles of operation through which the mechanism has moved.³

The hysteresis loop for a measuring instrument exhibits many of the characteristics of that of a structural material under stress, in that the loop is the narrower the less the range of operation (corresponding, in the case of the stressed specimen, to the range of stress); and in that it is the narrower the more delicate and workmanlike the construction, somewhat as the hysteresis loop of a stressed specimen is reduced in area with homogeneity and fineness of structure.

This curve, showing the readings (or errors in the readings) of the instrument over its whole scale range, plotted against corresponding values of the quantity being measured, for increasing and decreasing values of that quantity, is a valuable and in fact indispensable index to the operating characteristics of an instrument and affords distinctive and easily interpreted information regarding defects of design and workmanship discoverable with certainty in no other way.

³ Much of the discussion of this paper relates directly only to mechanical instruments whose indications are of a reversible character; of another class are time-measuring instruments and some types of integrating instruments in which the readings can not be made to repeat or recur at will; these are not completely amenable to the present treatment, but require special consideration. That this distinction is fundamental is clear from the fact that a chronometer can not be said to have sensitivity, since time is in its very essence epochal and irreversible. These methods do, however, apply with complete validity to many devices and mechanisms not always considered to be comprised in the term "instrument," e. g., telephone receivers and transmitters and phonograph reproducers, and a large class of controlling or value-limiting mechanisms, such as thermostats, barostats, hygrometers, voltage regulators, engine governors, and carbureters for internal-combustion engines, all of which are in fact in large measure subject to perturbations arising from effects to be classed as irreversibility. This is true, moreover, of liquid-column instruments such as mercurial barometers, for even in such instruments resistance to motion does not actually vanish at zero velocity of the liquid. Imperfect reversibility in measurable amount is known to exist in mercurial barometers and other liquid manometers.

Consider the curve shown in Fig. 1, which is plotted from observations on an automatic or self-indicating weighing scale. This particular scale is one in which the load is equilibrated by

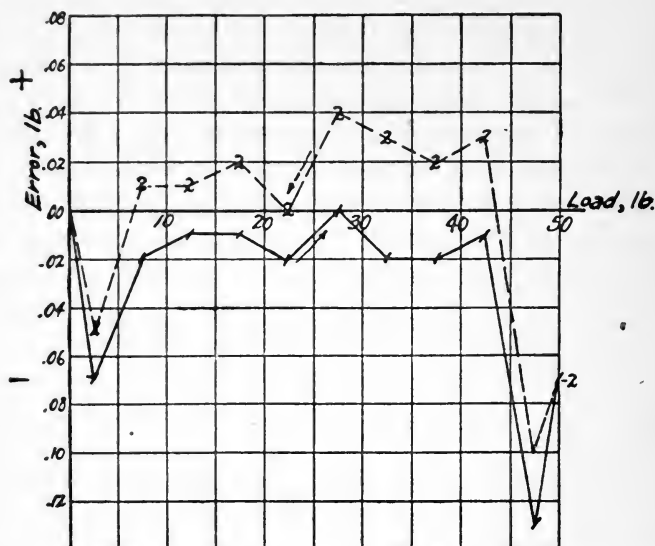


FIG. 1.—Hysteresis loop of automatic weighing scale of the cam-pendulum type (comprising no elastic force-measuring elements)

Note the manner in which the curve of decreasing readings reflects the aberrations of the curve of increasing readings, with a distinct tendency toward wider separation at the middle of the load range, so that the hysteresis loop would have a distinctly lenticular form, if the median line or mean error curve were rectilinear.

the variable turning moment of a pair of oppositely rotating pendulums mounted on ball bearings. The effect shown, therefore, is certainly not the hysteresis of inelasticity, such as would

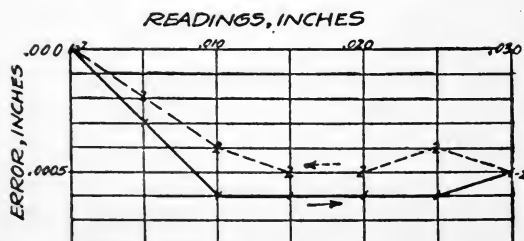


FIG. 2.—Hysteresis loop of dial indicator of the screw-train type, having high multiplication, an instrument much used for direct reading of small displacements and other linear magnitudes

be observed in the test of a spring scale, although, admittedly, elastic hysteresis is present in very small and negligible amount in every instrument, due to variations in the working stresses (in this case, in the load-sustaining parts of the scale which

transmit the forces from the platform to the pendulum). The sources of the hysteresis exhibited in the figure are manifold; though diverse in apparent character, however, all are expressible as backlash or are closely analogous to it.

Figs. 2 and 3 show other typical examples of hysteresis effects, the instrument of Fig. 2 having no source of hysteresis of inelasticity, while that of Fig. 3 does.

In one case so important was the recording of the backlash characteristic that the writer, in the inspection on a contract comprising 1200 automatic scales, found it advisable to note for each scale complete sets of readings taken at suitable intervals over the whole range of graduation for increasing and decreasing

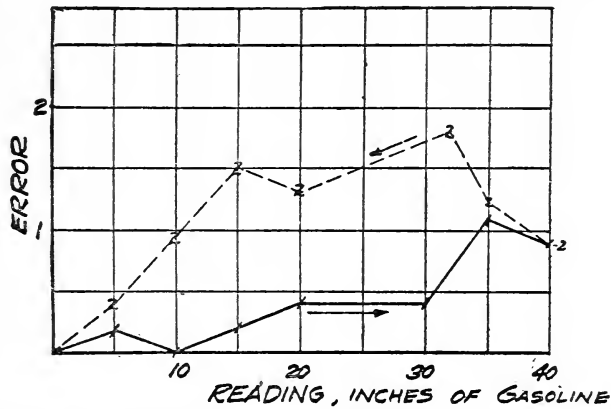


FIG. 3.—Hysteresis loop of depth gage for gasoline tank, a diaphragm type of pressure gage graduated in inches of gasoline. (Ordinates in inches of gasoline)

loads, the results of the test observations being plotted directly into graphs like that of Fig. 1. The median line of the loop so obtained gave the characteristic error of the scale, correctable by suitable alteration of the weights of the pendulums and of the contour of the cams upon which were wrapped thin metallic tapes transmitting the load from the platform system to the pendulums. The width of the hysteresis loop, moreover, gave a good index of the workmanship of pendulum bearings and connections.

It will readily be appreciated how valuable is the information which such a graph affords in evaluating the fundamental characteristics of an instrument and in indicating defects in workmanship and design.

2. TYPES OF VARIANCE

In all instruments which use a spring element for equilibrating the changes in the measured quantity directly or through conversion from displacements to forces, a source of variation exists in the imperfect elasticity of such a spring, whose changes in shape or dimensions are always to some extent irreversible. This imperfect reversibility of the stress-strain relation implies a hysteresis effect in the calibration of the instrument. Apart from this, we have the effects of imperfections in the mechanism itself, next to be discussed, as evidenced in such parts as journals, bearings, and gearing.⁴

(a) *Backlash*.—The function of turning pairs in a measuring instrument is distinctly different from that which they fulfill in a machine. In a machine the exact maintenance of the distances between centers of rotation of its elements is not ordinarily a fundamental requisite. Except in rare instances, only approximate invariability of distance relationships in the linkwork is required. In a measuring instrument, however, the whole result to be obtained depends upon the maintenance of constant or at least determinate intervals between the parts of the mechanism which transfer the forces or motions involved from the point of their reception to the point of registration or indication. This, expressed kinematically, requires that links must have a constant virtual length (if comprising only lower pairs) or determinate virtual length, constant and definite for any given configuration of the linkwork (if comprising higher pairs).

It will readily be seen that the presence of slack or backlash in the mechanism of an instrument will have the same effect on the indication as an equivalent ~~advent~~ displacement of those parts of the instrument receiving the force or displacement. The term backlash is here used to imply only looseness of fit resulting in play of the coacting parts. Such irregular and uncontrolled defects in the linkwork itself are usually, in fact, magnified at the point of registration, owing to the multiplication of the motion from the receiving to the registering element.

The effect of backlash is usually assumed to be that shown in the hysteresis loop at *B* in Fig. 4, viz, a rhombus of which the two

⁴ There are, of course, other sources of lag that are evidenced as loops in the calibration curves of instruments. As typical examples may be mentioned the irreversibility in the thermal expansion of some materials, and effects more or less closely a function of time, such as temperature variations occasioned by changes in the state of stress, lag in the drainage of liquid from the walls of a tube, and the thermal lag noticeable in certain types of electrical instruments and resulting from temperature changes accompanying the flow of current.

parallel horizontal sides (magnified, it may be, by the multiplication of the mechanism) correspond exactly with the geometrical clearance in the bearing. The portion *ab* represents the phase of load addition during which no change of indication occurs, pending the taking up of the lost motion. *cd* is the corresponding phase at the beginning of the returning portion of the curve. During this phase the pin or journal is assumed to be returning to contact with the opposite contact face of the bearing.

As a matter of fact, the only defect of workmanship which could produce this type of error is typified in the view at the left in Fig. 4, in which a round pin or journal coacts with an oval bearing, the lesser diameter of the bearing being exactly equal to the diameter of the journal. In this case the clearance of the journal in its

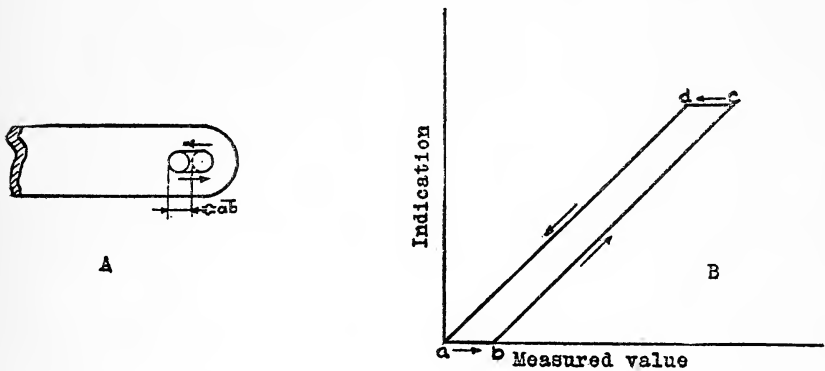


FIG. 4.—The rhomboidal type of hysteresis loop, rarely encountered in instruments

The unusual type of action of journal and bearing shown at A, or of toothed gears mating with circumferential clearance, will produce this form of calibration curve

bearing occurs in one direction only. Such a bearing detail could occur but rarely and then only in instruments of which the bearings were badly worn; moreover, a hysteresis loop of the form shown in B could hardly occur in any rationally designed instrument since some portion of the mechanism is usually restrained by a counterweight or spring (often a hairspring at the indicator spindle) to the end of keeping the journals in contact with the same general faces of the bearing, thus providing what is known as force-closure of the linkwork. Such a hairspring can not, however, eliminate all backlash effects so long as a finite "running" clearance exists between the components of the turning pairs, as will be shown hereafter.

Were an instrument to comprise only the form of backlash just defined, a constant displacement would exist between the upward

and downward branches of its hysteresis loops, so that a fairly definite and useful correction could be applied for the backlash error, if only the phase of the instrument movement were known.

As a matter of fact, the shape of the hysteresis loop for an instrument is lenticular, and not dissimilar to that of an imperfectly elastic structural specimen under stress. (To be sure, such a specimen with an applied extensometer is itself a measuring instrument, being a spring-controlled weighing scale with, ordinarily, a stiff spring and high magnification.) The reason for this shape of the hysteresis loop will be noted by reference to Fig. 5.

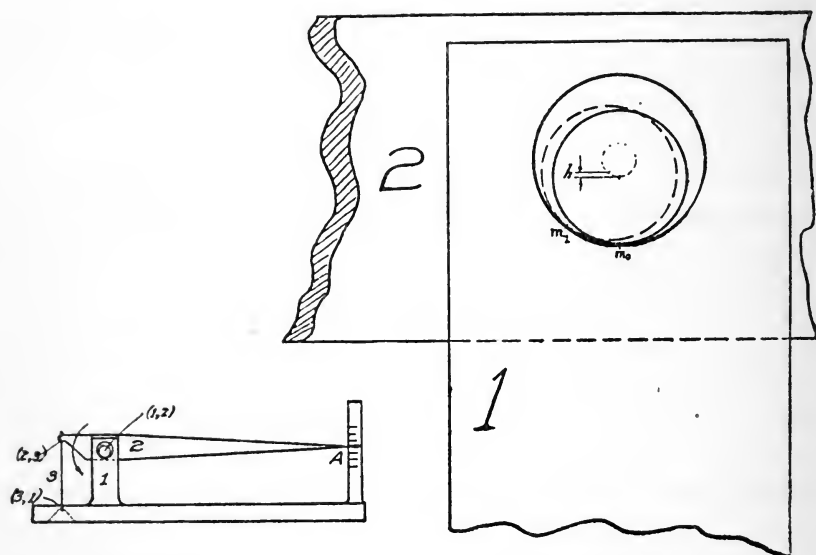


FIG. 5.—Illustrating the operation of instrumental "backlash" in introducing aberrations in the geometrical relation of journal to bearing.

Let 1 and 2 be two conjoined links of the mechanism of a measuring instrument. For the sake of concreteness, we may consider them comprised in the simplified hair hygrometer shown in the diagram at the left of the figure. During the operation of the hygrometer the links 1 and 2 rotate relatively to each other about some more or less definite center $(1, 2)$. The kinematic consideration of such a jointure would, of course, be to assume perfectly circular rotation about coincident geometrical centers of the journal and bearing. For most kinematic problems this approximation is amply close. In instruments, however, as has already been pointed out, the whole performance of the device depends upon the definiteness of the geometrical configuration at a given value

of the quantity being measured. Let us therefore examine into the action taking place in a single bearing having a loose journal. As 2 rotates relatively to 1—assuming that force-closure exists to keep the bearing and journal in contact at some point, and the hair 3 taut, as would be the case in any well-designed instrument—the journal in part rolls and in part slides around the interior of the bearing, its center describing the arc of a circle whose radius is equal to the difference in the radii of the journal and the bearing. The point $(1, 2)$ then for ordinary motions of the instrument will have a motion in this arc, which motion is superposed on the pure self-rotation which it derives from the intended operation of the mechanism. It is seen, therefore, that the motion of the pointer tip over the graduated reading scale does not bear a simple geometric relation to the lengthening or shortening of the link 3, being in fact distorted by the complex motion at $(1, 2)$.

In the case of a horizontal spindle turning in ball bearings, the exact performance may be somewhat different from the above, although the effect is the same. In cases which have been examined, the ring of balls surrounding the spindle appears to roll as a unit with the turning of the spindle, while the spindle itself remains in contact with, and has its weight supported by, one pair of balls near the bottom of each ball race. As the ring of balls turns, due to the rolling of the individual balls in the race, the spindle is carried upward an amount depending roughly upon the radial clearance between the spindle and the interior of the ring of balls.

The effect of this divagation of the point $(1, 2)$ in introducing error in the instrument reading may be simply expressed as the variation which it produces in the leverage ratio (in this case, in the two arms of the link 2). As the pin shifts to the left and upward in its bearing, a point which is the instantaneous center of relative motion of the two links and which lies at a determinate point situate between the center of the bearing in 1 and the contact point m_1 , shifts in such a way as to decrease the ratio of the two arms $\frac{(3, 2) - (1, 2)}{(1, 2) - A}$. On this account point A moves increasingly faster in relation to the point $(3, 2)$. When link 2 rotates in the opposite direction, the contact point rises along the right-hand face of the bearing, increasing the ratio $\frac{(3, 2) - (1, 2)}{(1, 2) - A}$, so that A will move increasingly slower in relation to the point $(3, 2)$. This mode of action will clearly account for the general lenticular

form of calibration curve exhibited by an instrument having backlash. In the case of instruments comprising spring force-resisting elements, such as occur in pressure gages, heat-engine indicators, aneroids, and the like, a similar loop will arise from inelastic action of the spring. This effect will be combined in the calibration curve with that due to backlash.

The limiting amount by which the point of contact may be displaced from the lowermost or most stable point of the bearing could, if necessary, be approximated from a knowledge of the weights of the parts, the direction and magnitude of the other forces involved, and the coefficient of friction of the materials of the journal and the bearing.

When the maximum sidewise displacement of the center of the journal is reached, a more or less steady⁵ state of relative slipping between journal and bearing takes place, so that insofar as concerns hysteresis from journal action, it is to be expected that the backlash loop corresponding to this portion of the operating range of the instrument will be flat. The loop will, therefore, be made up of three phases: A lower lenticular portion; a middle oblong, approximately straight-sided portion; and an upper lenticular portion. This state of affairs has been approximately confirmed by experiment. It may easily be that the calibration curve will not form a completely closed loop since local roughness and variations in surface conditions of bearing and journal may operate to prevent complete reversion of the parts to their initial positions. Specifically, this may occur when the force designed to effect force-closure of the system is insufficient in amount. The amount by which the loop fails of closure may be termed the set. The relative or specific set is the ratio of this residual deflection divided by the deflection which occasioned it.

In cases in which a link typified by 3, functions as a means of transferring or modifying a motion from 2 to 4, the effect of the interior rolling of journals in bearings is to lengthen link 2 and shorten link 4 for rotation of 2 and 4 in one direction, and vice versa for rotation in the other direction, thus again acting to widen the hysteresis loop.

Clearance between the engaging teeth of gears and toothed racks which appear in many instruments introduces backlash

⁵ It may even be that a step-by-step motion of the linkwork and indicator, due to the discontinuities in the relative motion, referred to above, can be distinctly observed, typifying in a sense the point-by-point manifestation of passiveness in preventing continuous response of the instrument mechanism; and reflecting the discontinuity between the values of the static and kinetic friction of the linkwork.

effects of the same general nature as those outlined above. In many cases the backlash effect in a gear train will be superposed upon that of the journals and bearings of the train. In every case there are certain phases of motion in which such effects will be additive. In gear trains, moreover, there is especial likelihood of occurrence of the true backlash type of loop shown in Fig. 4 at *B*, since clearance at the pitch line will tend to result in actual discontinuity in the transference of motion from one part of the train to another, whenever the direction of rotation is reversed.

(b) *Irregular Variance*.—In the case of instruments which are characterized by poor workmanship or are in a state of ill repair, the hysteresis loops obtained on successive runs may be far from concordant in either shape or magnitude, this condition being of course a result of inaccurate fitting of such serious order that variations in the friction and journal displacement, even for a par-

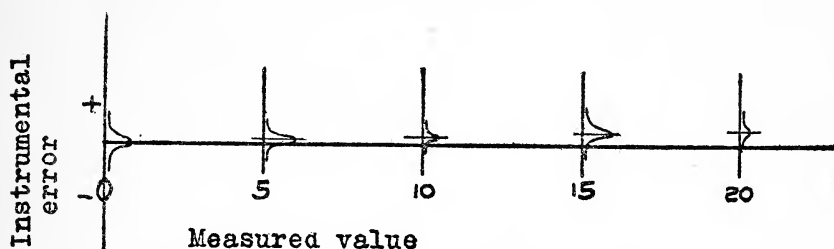


FIG. 6.—Illustrating the representation of irregular variance by reference to a family of probability curves developed at various points along the instrumental scale

ticular indication, become significant. In the case of instruments of this kind, the variance may be well expressed and defined by plotting frequency curves of the readings or errors obtained for a given value (or for a series of definite values) of the measured quantity repeatedly applied.

One series of points will be plotted for readings taken at varying rates and amounts of increase of the measured quantity terminating in the value corresponding to the particular point of the scale under investigation, and another set for decreasing values terminating at the same point. Such a curve, in which frequency of occurrence of a particular reading or error is plotted against the true value of the measured quantity at that reading, gives the probability of occurrence of any amount of variation from, say, the mean instrumental reading.

A series of such frequency curves taken at various parts of the reading range may be plotted as shown in Fig. 6, in which case

the surface which may be conceived of as enveloping the family of probability curves when they are thrown into three dimensions, will enable one to determine the probability of a given error at any point along the reading scale. The principal utility of such a graph, however, will be in the criterion which it affords as to the comparative performance of instruments of diverse design or workmanship.

Lest it be felt that the importance of the hysteresis error due to bearing clearance and backlash has in the foregoing been overestimated, the following should be given consideration. It may well be true that in some instruments, especially those of the least mechanical complexity, the error is small; in other cases it may have an appreciable value and yet be of minor importance as compared with the hysteresis due to imperfect elasticity of elastic resistor elements and as compared with the drift or elastic after-effects, which also frequently accompany spring force-resisting elements. Nevertheless, in many instruments the backlash type of error arising in purely mechanistic sources is the one which governs the practicability of the desired accuracy and sensitivity. As examples, automatic or self-indicating weighing scales and balances, especially those using the principle of pendulum inclination, may again be called to mind. Another important class of instruments, in which the backlash effect is the limiting and preponderant factor, is that of length-measuring instruments commonly employed for measuring small differences of displacement. Examples of these are found in the dial-and-plunger instruments which use wheel or screw trains of high multiplication to convert rectilinear displacements at the receiving point into rotations of an indicator over a graduated dial. Were it not for the variance effects of mechanical hysteresis in these instruments they would obviously be applicable to the most refined length comparisons. In point of fact, one of the types most used, due to unavoidable deficiencies in the fit of the wheelwork spindles and meshing gears, shows variance differences amounting to several graduations, at a given displacement of the receiving point, the dial of this instrument being divided to $\frac{1}{10\,000}$ inch of plunger displacement. This, then, is an example of an instrument in which variance errors have been the limiting factor; the extremely high magnification and considerable range of reading, with the necessarily large number of coacting parts required thereby, makes their minimization a very difficult matter.

Almost without exception, the complete calibration of measuring instruments of nearly every type, not excepting those using a liquid meniscus as indicating element, should include a determination of the variance characteristics. The magnitude of the variation shown in this determination will indicate whether or not it will be worth while to analyze the hysteresis effect into the factors which compose it, viz, the backlash effect, the imperfect elasticity effect, and the drift or elastic aftereffect.

(c) *Drift*.—Another phenomenon of considerable importance in some types of instrument is that of drift, which is a time effect characterized by a more or less gradual movement of the indicator, asymptotically to a definite reading, after all conditions external to the instrument have become constant. It would seem that very little of this effect can arise from causes originating in the kinematic relations of the linkwork; it must be due chiefly to elastic aftereffect or the hysteresis of inelasticity in stressed portions of the measuring elements. If we accept, as applying to stresses other than torsional, the conclusions of J. J. Guest and F. C. Lea in their recent paper⁶ on hysteresis phenomena, we may ascribe all such effects to conditions of actual overstrain in imperfectly elastic force-measuring or force-sustaining parts. An example is seen in the aneroid barometer, where drift takes place as a result of slow settling of the diaphragm or spring to a steady state of stress. It is quite reasonable to suppose that in corrugated diaphragms, such as are used for this purpose, zones of decided overstrain do exist.

3. MEANS FOR REDUCING VARIANCE

(a) *Details of Design*.—Several possible remedies are to be considered for minimizing the hysteresis loop of measuring instruments arising from the causes described. The most obvious one is, of course, to improve the fit of journals in their bearings to such an extent that the bearing clearance is negligible. This procedure is oftentimes mechanically impracticable for several reasons: First, that when an ideally close fit is had, the journals and their bearings must have cross sections which are perfectly circular, since otherwise any relative rotation would result in binding; second, slight differences in the temperature of the parts or in the condition of the lubricant, or the introduction of dust in the use of the instrument, would increase the friction to an

⁶ J. J. Guest and F. C. Lea, Torsional Hysteresis of Mild Steel, Proc. Royal Soc., June 1, 1917.

amount fatal to proper operation. Perhaps the best type of bearing to reduce these difficulties is the conical pivot, consisting of two oppositely directed cones engaging in suitably supported conical hollows, both the cone and its mating depression being quite sharp, and suited to each other as to included angle. This requires that the conical depression shall include a somewhat larger angle than the cone, in order that complete contact occur only at a point or infinitesimal surface in the axis of rotation. This is the type of bearing commonly used for the balance wheels of alarm clocks and low-priced watches. While having relatively low friction and maintaining a practically invariable relation of bearing to journal, it can not, of course, withstand any considerable load.

Another method, which looks very promising in the limited application it has had in weighing scales and a very few other

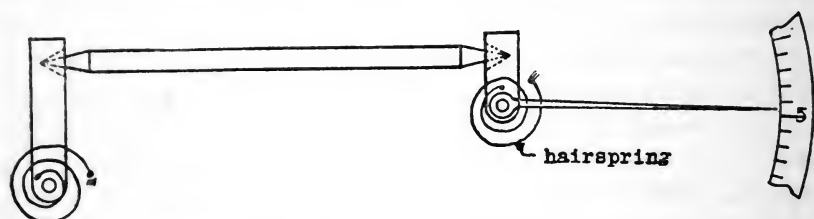


FIG. 7.—A simple type of connector having low frictional resistance and affording satisfactory constancy in distance relationships

measuring instruments, is the use of a flexible or ribbon-like elastic connector, secured to the conjoined members by means of clamps or equivalent device. With such a connector, the constancy of distance between centers of relative rotation for a given configuration of the system is assured in a practically perfect manner, neglecting temperature and similar effects which are common to all types of bearings. The only contribution of this type of connector to the total hysteresis of the instrument lies in its own elastic hysteresis, which should be very small indeed for very thin bands or ribbons of steel or similarly favorable material, operated as they commonly are, through but small or moderate arcs of motion.

Knife-edges in V-shaped bearings or single cone points and cups arranged as shown in Fig. 7 may be used if the shape and position of the bearing and edges are such as to insure the contact remaining always at the intended point. It is difficult to see, however, where this arrangement can offer any important advan-

tage over the flexible connector just described, except in the case of precision instruments for use in the laboratory, where the most careful operation and handling can be assured.

The flexible connector has another important property which has been utilized to a considerable extent in the design of automatic scales, viz, that a correction of the motion of the parts to obtain a linear or uniform scale of graduations can readily be secured by arranging that one or both ends of the tape wind upon a cam of suitable contour, a feature whose importance will readily be appreciated. It does not seem that the use of the flexible connector in such instruments as pressure gages (including aneroids), recording thermometers, hygrometers, tachometers, and the like has been as extensive as the favorable properties of the device warrant, and a marked improvement in the constancy or reproducibility of reading of such instruments should be secured by employment of this arrangement. Critical examination of the usual pressure gages and tachometers, as well as complex recording instruments generally, will convince one that the importance of special care in the design of the turning pairs in the linkwork has not been commonly appreciated, and it does not seem that the best practicable performance of such mechanisms has been approximated.

In connection with the pressure-measuring element familiar in the aneroid, an important development has been noted in certain recording instruments, and is now being manufactured for use in aviation instruments of the indicating type, its object being to eliminate gearing and jointed linkwork. Instead of the common diaphragm or monoplanar Bourdon tube of small movement, a helically coiled Bourdon tube is used, connected at one end to the fixed frame of the instrument, and at the other, through a temperature compensator which need not be described, to the indicator spindle, which latter is concentric with the helix. As pressure is applied, the Bourdon tube tends to uncoil, and drives the coaxial pointer directly, without the need for gear or link trains of any kind. A considerable arc of pointer movement can easily be obtained in this way without complication, and it would seem, with very satisfactory reliability. The importance of improvement in this field is appreciated when it is known how extensively the compact and inexpensive pressure-diaphragm construction has recently been applied in the development of every variety of aviation instrument, in which use, of course, the direct, pointer-reading type is to all intents and purposes a necessity as against

the liquid-column manometer. The diaphragm capsule and the Bourdon tube are now employed in thermometers, air-speed indicators, altimeters, liquid-depth gauges, and in many other applications apart from the simple determination of static pressure.

(b) *Effects of Vibration.*—With the considerations developed in the foregoing in mind, it is easy to see how the effects of vibration treatment suitably applied to an instrument at the time of reading will tend to correct the errors introduced by friction and backlash. It is a well-known fact that with certain instruments, such as Bourdon tube or diaphragm pressure gauges, the reading tends to settle toward a more reliable and reproducible value if the instrument is tapped or jarred. The interaction of parts which brings this about will be perfectly in line with the facts developed above. Briefly the effect is of two related sorts: (1) The minimization of static friction by the momentary disengagement or separation of coacting bearings and journals, and (2) the opportunity which the same disengagement affords, for journals to settle to the "minimum" or base point of contact in their bearings. On the basis of this reasoning it would seem that the best results from vibration treatment before a reading is taken will ensue when the vibration starts with a relatively large amplitude and decreases finally to practically zero amplitude, since the slight shifting about the optimum indication, due to the amplitude of the vibration itself, is thus eliminated. In brief, the effect of vibration is to permit the forces which tend toward a definite, unique, equilibrium point and are inherent within the instrument, to operate against the minimum possible resistances in moving the parts toward that equilibrium; in another sense, energy is applied from without the system to replace that unavoidably lost in the deficiencies of the mechanism. On this point an analogy with magnetic hysteresis exists. A substance is most accurately brought to a given state of magnetization by subjecting it to a field whose intensity oscillates with decreasing amplitude about that (or corresponding) value of magnetization.

A priori, it would seem that by vibration treatment judiciously applied at the time a reading is taken, errors of result due to the mechanical sources of variance can be almost completely eliminated.

4. RELATION OF VARIANCE TO USEFUL SENSITIVITY

It is often found that particular measuring instruments are given a sensitivity far higher than warranted in the face of the error obtainable in reading and resulting from the variance present.

Similarly, the graduation of instruments is often found to be far closer than the large amount of the variance justifies. Care should be taken in the design of measuring instruments that the units of graduation and the openness of the scale are not out of all proportion to the effective reproducibility of reading possible. For testing or laboratory instruments, the mean interval of graduation should not be less than five times the mean variance, since it is to be expected that observations will be noted accurately to one-fifth the smallest graduation or less. For commercial instruments the unit of graduation and the variance may be more nearly equal, say in the ratio of 2 to 1.

It is obviously misleading and absurd to graduate a tachometer to a single mile per hour or revolution per minute when the reading at a given true speed may vary as much as 5 miles per hour or revolutions per minute. Such inconsistencies, however, are very frequent in practice, as is the related but less serious one of having a needlessly open scale on a very variant instrument.

Unnecessary time and care in taking readings are required when the sensitivity is ill-proportioned to the variance. Instruments are not common which possess such refinement of workmanship that the characteristics of successive hysteresis loops are uniform enough to permit of practical correction for the variance of observed readings. For these reasons the sensitivity may easily be disadvantageously high, in tending to induce an erroneous estimate of the precision of results, and if this mistake does not occur, in requiring careful investigation of the particular instrument involved to determine the portion of its maximum sensitivity which is really and practically available in service.

Other factors may well have a part in determining the sensitivity to be aimed at in a given instance, for example: In an equal arm balance it is preferable to use a relatively low inherent sensitivity, enhancing the visibility of changes in the deflection of the balance beam by the use of an optical lever, microscope and scale, or similar external means. In this particular case the advantages gained are (1) decreased period of oscillation, which reduces the drift and other errors flowing from the protraction of the observations over a long period of time; and (2) the greater ease of manipulation and control of a balance having low inherent sensitivity.

The factors of inaccuracy (or accuracy), sensitivity, specific set and variance may, as a convenient means of arriving at and expressing the "figure of merit" of an instrument, be referred to

the total range of graduation instead of to the value of the measured quantity under observation, as presented in the definitions of the foregoing pages. For example, it will be of service to express the maximum set or maximum variance observed in the reading of an instrument in terms of its ratio to the total range of values represented in the graduation or use of the instrument, in order to arrive at a single significant number representing a measure of utility or merit of the instrument with regard to the particular property in question. In like manner the reciprocal of any of these quantities (that is, range of graduation divided by maximum set or maximum variance, etc.) may be similarly employed.

V. SUMMARY

The terms accuracy, sensitiveness, and variance in reference to the characteristics of a measuring instrument are most useful when quantitatively defined. In practice the first is best expressed in terms of its reciprocal, as the ratio of error, arising from whatever cause, observed in an indication of the instrument, divided by the true value of the quantity being measured, this value being the *specific inaccuracy* at that reading. This factor is useful in rating the performance of an instrument, while the absolute error observed is used as a means of correcting observations for use in calculation.

Sensitivity is the rate of displacement of the indicating element with respect to change of the measured quantity. The mode of expressing sensitivity should be definite and not dependent upon the observer's personal judgment; moreover, when possible it should be so chosen as to be independent of the accident of dimensions in the indicating element where such dimensions or the final magnification of indicator movement are within reasonable limits alterable at will.

The effect of passiveness or sluggishness in the action of instruments is to be clearly distinguished from insensitiveness. Emphasis is laid upon the fact that sensitivity can not be determined by direct measurement unless friction and lost motion are sensibly eliminated or are negligible. The amount of the passiveness at any point of the instrumental scale is measured by the smallest alteration in the quantity to be measured which will produce any change whatever in the indication of the instrument. The relative sluggishness or *passivity* at any point is that change of value of the measured quantity which effects the first perceptible re-

sponse in the indication of the instrument, divided by the initial value of the measured quantity.

The term variance is to express all changes of indication intrinsic within the instrument itself, and not per se indicative of change in the measured quantity. It is defined as the range, at any given value of the measured quantity, of variation in reading which may be exhibited by the instrument under repeated application of the same value of the quantity being measured, after a steady reading has been attained, the environment remaining unchanged. The specific variance or *variancy* is the ratio of the range of variation above defined divided by the value of the measured quantity itself.

Variance arises in three causes of mechanistic character: (1) Backlash, the operation of which in producing variance effects is fully discussed in the complete paper, and (2) friction, the primary manifestation of the latter being passiveness, discussed above. Moreover, (3) in all instruments using any of the various forms of springs as the force-resisting or restoring element, variance arises in the imperfect elasticity of that spring; this source of instrumental hysteresis is a direct reflection of the hysteresis loop in the stress-strain relation of the material composing the spring. The variance, which embodies all of the foregoing sources of variation from unique reading for a given value of the quantity to be measured, represents the range of uncertainty of indication exclusive of the factor of drift, which, being a time effect, demands special consideration. The latter effect, however, is peculiar to instruments in which a considerable range of stress in the parts occurs during their operation, or in which the elastic properties of the elements of the mechanism are unfavorable. The set of an instrument is the amount by which the indicator fails to return to its initial position after a deflection has occurred. The relative or *specific set* is the ratio of this residual deflection divided by the deflection which occasioned it.

The errors manifested as variance are often of great importance and should be examined into before they are assumed to be negligible, as they occur in all types of instruments and often have the effect of delimiting the field of useful application of the instruments.

The variance of an instrument of good characteristics can be defined by reference to the hysteresis loop obtained through cyclic variation of the measured quantity, against the values of which quantity are plotted the corresponding readings (or errors) of the

instrument. In other cases the variance can be expressed by reference to a family of probability curves giving the frequency of occurrence of each particular reading (or error) for a given value of the measured quantity (or, conversely, giving the frequency of particular deviations of the measured quantity for a given reading).

Calibration curves of typical instruments are presented to show the character of the hysteresis loops and the nature and amount of variance errors as actually determined.

Instrumental variance may be reduced by simple changes in details or simplification of design, the former relating especially to the points of jointure in the link work. Important factors in the design of such connections have failed of recognition in the development of many common instruments. Vibration of some kinds of instruments at the time the reading is taken will reduce the variance on account of its effect in minimizing the static friction opposing the motion of the parts of the mechanism to the position of equilibrium.

The amount of variance determined for an instrument should establish the optimum sensitivity to be sought in its adjustment, and also, roughly, the scale interval. High sensitivity or needlessly minute graduation, when accompanied by high variance, are likely to be both uneconomical and misleading. A working basis for determining the sensitiveness to be sought in the adjustment of the instrument, and the closeness of graduation can readily be established.⁷

WASHINGTON, February 9, 1918.

⁷ Apr. 15, 1919. Other papers by the present author discussing the subject of instrument variability, which have appeared since the foregoing was printed as a separate, are: *The Determinateness of the Hysteresis of Indicating Instruments* (Journ. Wash. Acad., Vol. 9, No. 2, Jan. 19, 1919) and *The Concept of Resilience with Respect to Indicating Instruments* (Journ. Frank. Inst., February, 1919).

The principal modifications and extensions of the treatment of the present paper are with respect to the exact specification of the conditions under which hysteresis loops in instrument calibration are to be taken in order to obtain concordant results; the demonstration of the high order of reproducibility of the hysteresis loops observed under such conditions; and the principle of rating an indicating instrument on the basis of the smallness of area of the hysteresis loop corresponding to a specified range of operation, which affords a numerical measure of the perfection of energy restoration or resiliency of the instrument, and, hence, of the invariability or consistency of its indications when it is applied to unregularized, acyclic use.

On page 747 after the word "for" in the third line of the last paragraph, understand the words "slowly and aperiodically." For reasons, consult second reference above. Referring to page 755, "(b) Irregular Variance," this method now appears particularly applicable to integrating instruments, since indicating instruments even of the crudest sorts, recent results show, can be so operated as to give very concordant successive calibrations. The probability curves of Fig. 6, instead of being slowly asymptotic to the several vertical axes, should meet those axes a short and definite distance from the respective normals of the curves.







